



Gil Gonalo Freire Martins

Um novo modelo h brido multivariado entre uma Rede Neuronal Artificial e um modelo ARIMA para a previs o de taxas de c mbio

Trabalho de Projecto

Janeiro de 2014



UNIVERSIDADE DE COIMBRA



FEUC FACULDADE DE ECONOMIA
UNIVERSIDADE DE COIMBRA

Gil Gonalo Freire Martins

Um novo modelo h brido multivariado entre uma Rede Neuronal Artificial e um modelo ARIMA para a previs o de taxas de c mbio

Disserta o de Mestrado em Economia, apresentada   Faculdade de
Economia da Universidade de Coimbra para a obten o do grau de Mestre

Orientador: Professor Doutor Pedro Ba o

Coimbra, 2014

Numa sala lotada, só necessita de ver uma polegada acima dos
restantes para perceber coisas que todos os outros irão perder

James Derrick Slater

Resumo

Apesar dos muitos modelos propostos, é notória a dificuldade na previsão de taxas de câmbio. Estudos recentes com redes neurais artificiais (RNA) sugerem que estas podem ser uma alternativa a outros modelos para a realização de previsões, quando existem não linearidades nas séries temporais. Este estudo propõe uma abordagem multivariada, aplicando um modelo híbrido entre uma RNA e um ARIMA, de forma a aumentar a capacidade de previsão da RNA, dado que uma série temporal pode possuir componente não linear e linear. Implementando primeiro um algoritmo genético adaptativo para construir a estrutura da RNA, são consideradas duas taxas de câmbio mensais e semanais do Euro e do Yen contra o Dólar, para inferir sobre a qualidade do modelo híbrido multivariado proposto face a outros modelos. Os resultados empíricos das previsões fora da amostra, indicam que para as séries mensais o modelo híbrido não é vantajoso, devido à não existência de não linearidades nestas. No entanto para séries semanais, obtêm-se vantagens da utilização do modelo híbrido. Assim uma abordagem híbrida multivariada pode ser pertinente para a previsão taxas de câmbio semanais e possivelmente diárias, ao invés da utilização dos modelos individuais ou híbridos univariados.

Palavras-chave: Previsão de Taxas de Câmbio, Redes Neurais Artificiais, Modelo ARIMA, Modelo Híbrido, Algoritmo Genético.

Classificação JEL: C45, C53, F31

Abstract

Despite the many proposed models, it is notoriously difficult to predict exchange rates. Recent studies with Artificial Neural Networks (ANN) suggest that these may be an alternative to other models for conducting forecasts when there are nonlinearities in the time series. This study proposes a multivariate approach, applying a hybrid model between an ANN and an ARIMA in order to increase the predictive power of the ANN, since a time series may be composed of linear and non-linear components. Implementing first a adaptive genetic algorithm to build the structure of the ANN. Two monthly and weekly rates of the Euro and the Yen against the Dollar are considered to infer the quality of the proposed hybrid multivariate model compared to other models. The empirical results of the out-of-sample forecasts indicate that for monthly series the hybrid model is not advantageous due to the non existence of nonlinearities in these series. However for weekly series, we obtain benefits of using the hybrid model. Therefore the multivariate hybrid approach may be pertinent for forecasting weekly and possibly daily exchange rates, instead of using individual models or univariate hybrid models.

Keywords: Forecasting Exchange Rates, Artificial Neural Networks, ARIMA model, Hybrid Model, Genetic Algorithm.

JEL Classification: C45, C53, F31

Sumário

1	Introdução.....	1
2	Revisão da Literatura.....	3
2.1	Teorias da taxa de câmbio	3
2.2	Redes Neurais Artificiais.....	5
3	Construção da base de dados.....	9
4	Metodologia: Modelos e Algoritmos.....	11
4.1	Rede Neuronal Artificial	11
4.2	Metodologia para construção de uma RNA	14
4.2.1	Algoritmo genético adaptativo proposto	14
4.2.2	Regularização no treino da RNA.....	17
4.3	Modelos de referência	18
4.3.1	Modelo ARIMA	18
4.3.2	Modelo híbrido convencional para séries temporais univariadas.....	18
4.3.3	Modelo de Passeio Aleatório	19
4.4	Testar a não linearidade: Testes BDS e LWG	19
4.5	O modelo híbrido multivariado proposto.....	20
5	Resultados empíricos.....	21
6	Conclusão	25
7	Bibliografia.....	26
8	Apêndices	29

1 Introdução

As taxas de câmbio estão entre os mais importantes índices económicos, sendo relevantes para os mercados financeiros, para decisões de política económica e para decisões de investimento em geral, sendo afetadas por múltiplos fatores económicos ou mesmo psicológicos. Há vários modelos que tentam explicar as flutuações das taxas de câmbio. No entanto, Meese e Rogoff (1983) demonstraram que um simples modelo de passeio aleatório não produz resultados piores que os modelos lineares macroeconómicos testados, e existentes até à data, para períodos até um ano. Pelo que concluíram que as taxas de câmbio seguem a teoria dos mercados eficientes. Cheung et al. (2005) repetiram o estudo de Meese e Rogoff, utilizando modelos mais recentes, desenvolvidos entre as datas dos estudos. Contudo, chegaram às mesmas conclusões que Meese e Rogoff.

A previsão de taxas de câmbio é assim um desafio com um grau elevado de dificuldade. Devido aos múltiplos fatores que as determinam, as taxas de câmbio apresentam um comportamento complexo, volátil e possivelmente não linear. Os métodos propostos para resolver este problema podem ser divididos em duas categorias: os métodos estatísticos, como os modelos paramétricos não lineares ARCH ou GARCH, e os métodos baseados em inteligência artificial, como redes neuronais artificiais ou máquinas de suporte de vetor. Este estudo focar-se-á de forma particular nas redes neuronais artificiais.

De facto, vários estudos demonstram que abordagens não lineares, multivariadas ou univariadas, baseadas em Redes Neuronais Artificiais (RNA) se comportam melhor que modelos lineares ou de passeio aleatório. Existem várias categorias de RNA. Contudo, neste estudo, quando se fala de RNA, fala-se de Redes Neuronais Artificiais Alimentadas para a Frente.

As RNA possuem diversas vantagens sobre os modelos lineares e outros modelos não lineares, entre as quais está o facto de conseguirem modelar até certo grau relações lineares entre as variáveis. Devido à sua natureza não-paramétrica, não é necessário possuírem *a priori* conhecimento sobre as relações entre as variáveis. Tornam-se assim menos suscetíveis a má especificação do modelo (Khashei & Bijari 2011a), o que lhes permite testar múltiplas relações entre as variáveis, pelo que são capazes, em teoria, de aproximar qualquer função contínua (Falát 2012). Esta propriedade pode revelar-se vantajosa para a previsão de taxas de câm-

bio, dada a dificuldade que tal encerra. Uma desvantagem é o facto de não ser possível interpretar os valores das estimações da RNA, ao contrário de outros modelos.

A existirem, todavia, serão raras as séries económicas completamente lineares ou não lineares. Assim, para melhorar o desempenho da previsão, é razoável assumir que estas são constituídas por uma componente linear e outra não linear (Zhang 2003b). Por conseguinte será vantajosa a hibridização de modelos não lineares com lineares.

Foram publicados vários estudos, para diversos tipos de séries temporais, no âmbito da hibridização de RNA com modelos lineares. No entanto, focam-se em abordagens univariadas, seguindo as metodologias propostas por Zhang (2003b) ou por Khashei e Bijari (2011a). No que se refere a uma hibridação entre uma RNA multivariada com modelos lineares, nenhum estudo, ao que se pôde apurar, foi publicado para qualquer série temporal.

Este estudo pretende assim colmatar esta lacuna na bibliografia, propondo um novo modelo multivariado de hibridação entre uma RNA e um modelo ARIMA. Utilizaremos taxas de câmbio para testar o novo modelo proposto e dar resposta às seguintes questões: Poderá a modelação dos resíduos da RNA multivariada por um modelo ARIMA aumentar a qualidade de previsão? São os mercados cambiais eficientes para amostras semanais e mensais? Existe vantagem em utilizar uma abordagem híbrida ou multivariada na previsão de taxas de câmbio?

Para encontrar respostas às questões acima referidas, este estudo utiliza o coeficiente da correlação da distância¹, a fim de seleccionar as potenciais variáveis de entrada (variáveis independentes) da RNA que revelam maior correlação com a variável de saída (variável dependente), e assim reduzir o número de potenciais entradas a considerar pela RNA. Posteriormente constrói-se o modelo híbrido multivariado entre uma RNA, seleccionada por um algoritmo genético, e um modelo ARIMA para a previsão um período à frente das taxas de câmbio mensais e semanais para os pares de moedas EUR/USD e JPY/USD.

Os resultados do novo modelo híbrido multivariado proposto são posteriormente comparados a um modelo de passeio aleatório que serve como referência. São ainda utilizados como comparação um modelo ARIMA, uma RNA multivariada e o modelo híbrido univariado proposto por Khashei & Bijari (2011a). A organização do estudo é a que se segue. Na secção 2 é discutida a bibliografia e estudos passados relevantes. Na secção 3 apresenta-se o processo de

¹ Do inglês *Distance Correlation*.

construção da base de dados. Depois, na secção 4, apresentam-se os modelos e as metodologias utilizadas neste estudo, analisando-se de seguida os resultados empíricos na secção 5. Finalmente conclui-se na secção 6.

2 Revisão da Literatura

2.1 Teorias da taxa de câmbio

Os fatores determinantes das taxas de câmbio podem ser vários, assim como as várias teorias que existem para a formação destas. Contudo, há duas teorias principais para explicar o comportamento das taxas de câmbio: a teoria da paridade das taxas de juro e a teoria da paridade do poder de compra. Estas duas teorias abordam a determinação da taxa de câmbio a partir de mercados diferentes. No caso da teoria da paridade das taxas de juro, o ponto de partida é o mercado financeiro internacional. No mercado financeiro, os agentes podem comprar e vender títulos denominados em diferentes moedas. A escolha das moedas de denominação dos títulos a comprar e a vender depende do rendimento proporcionado pelos títulos, o qual depende das taxas de juro nos diferentes países e da variação esperada da taxa de câmbio entre as moedas envolvidas. Se a taxa de juro no país de origem for i , a taxa de juro no país estrangeiro for i^* , a taxa de câmbio inicial for E_0 (unidades de moeda nacional por uma unidade de moeda estrangeira) e a taxa de câmbio esperada para o final do investimento for E_1 , então um investidor nacional, tendo de converter moeda nacional em moeda estrangeira para comprar títulos estrangeiros, comprará títulos estrangeiros se:

$$1 + i < (1 + i^*) \frac{E_1}{E_0} \quad (1)$$

Se esta situação for comum a vários investidores, a troca de moeda nacional por moeda nacional e a compra de títulos estrangeiros farão com que, por um lado, a moeda nacional, cuja oferta no mercado cambial aumentou, se deprecie no presente (aumento de E_0), e, por outro lado, a taxa de juro dos títulos estrangeiros, cuja procura aumentou, diminua (diminuição de i^*). Assim, num mercado concorrencial, a tendência será para a igualdade (se a desigualdade fosse a oposta, o raciocínio seria similar), ou seja:

$$1 + i = (1 + i^*) \frac{E_1}{E_0} \quad (2)$$

Uma aproximação linear a esta equação produz a equação usualmente associada à paridade da taxa de juro:

$$i \approx i^* + \Delta e \quad (3)$$

Nesta equação, Δe representa a taxa de depreciação da moeda nacional. Se tomarmos a taxa de câmbio como a variável endógena, esta equação diz que a taxa de câmbio é determinada pelas taxas de juro, nacional e estrangeira, e pela taxa de câmbio esperada. Ainda poderíamos acrescentar ao raciocínio que nos trouxe até esta equação, considerações acerca do risco envolvido no investimento. Dessa forma, concluiríamos que o grau de risco também influencia a taxa de câmbio. Igualmente, poderíamos, por exemplo, admitir que o mercado cambial não está sempre próximo do que seria um mercado concorrencial, nomeadamente quando um banco central, emissor de uma das moedas, intervém. Estas intervenções, ou o que quer que as determine, irão influenciar o comportamento da taxa de câmbio.

A teoria da paridade do poder de compra aborda a determinação da taxa de câmbio a partir do mercado de bens. Se os agentes puderem comprar o mesmo bem em dois países, irão preferir o que for mais barato. Novamente, na ausência de impedimentos à concorrência, o preço tenderá a ser igual nos dois países, quando convertido na mesma moeda. Se o preço nacional for P e o preço estrangeiro for P^* , teremos:

$$P = P^* E_0 \quad (4)$$

Tomando logaritmos, representados por letras minúsculas, passamos a ter:

$$p = p^* + e_0 \quad (5)$$

Tomando a primeira diferença, obtemos a versão relativa (em vez de absoluta) da teoria da paridade do poder de compra:

$$\Delta p = \Delta p^* + \Delta e_0 \quad (6)$$

Ou seja, a taxa de depreciação da moeda nacional depende da taxa de inflação nacional e da taxa de inflação estrangeira, ou daquilo que as determinar, como, por exemplo, a evolução da oferta de moeda em cada país (abordagem monetária).

2.2 Redes Neurais Artificias

Em face das vantagens das RNA, nomeadamente o facto de serem não-paramétricas e aproximadores universais de qualquer função continua, não é de espantar a atenção recebida por estas no que respeita à previsão de séries temporais. Alguns estudos foram elaborados utilizando RNA para a previsão de taxas de câmbio, apresentando resultados contraditórios, embora geralmente favoráveis às RNA.

No que respeita à taxa de câmbio EUR/USD, os trabalhos que utilizam RNA não rejeitam de forma conclusiva a hipótese de mercado eficiente (Dunis & Williams 2002; Seliem 2006; Pacelli 2011). Por sua vez, Lixia (2009) testa a existência de não linearidade em várias taxas de câmbio diárias, através do teste *BDS* e do método *Surrogate*, e conclui que apenas a EUR/USD apresenta comportamento aleatório; as restantes (USD/CAD, JPY/USD e USD/GBP) apresentavam características não lineares.

A utilização de redes neuronais com dados mensais, comparativamente a dados semanais ou diários, também parece apresentar uma baixa capacidade para a previsão de taxas de câmbio, comparativamente a um modelo de passeio aleatório (Seliem 2006). Contudo, esta conclusão pode estar dependente da dimensão da amostra; no texto mencionado, a dimensão da amostra não é referida.

De facto, as RNA necessitam de um maior número de observações do que, por exemplo, as 100 observações consideradas necessárias para obter bons resultados com um modelo ARIMA (Zhang 2003a). Já Hann e Steurer (1996) comparam a performance de diversas RNA contra modelos lineares para a taxa de câmbio USD/DEM e concluem que, ao contrário dos dados semanais, nos mensais as Redes Neuronais não têm um desempenho consideravelmente melhor que o modelo de passeio aleatório. Huang et al. (2006) concluem também que as RNA são mais apropriadas para modelar dados semanais que dados mensais, no caso, das taxas de câmbio USD/GBP e USD/JPY. Um problema destas conclusões, que se terá em conta neste trabalho, deve-se ao facto de estes estudos não terem tentado controlar de forma eficaz o problema do “sobre-ajustamento” dentro da amostra. Este problema resulta da utilização de relativamente poucas observações mensais para estimar um número elevado de parâmetros, como é geralmente o caso das RNA. A consequência é que o modelo pode ter um bom desempenho

na amostra para o qual foi (sobre-) ajustado, mas não conseguir representar devidamente o comportamento de outras amostras, ou seja, não ser generalizável.

Kaastra e Boyd (1996) discutem o facto de poucos terem conseguido modelar uma RNA de forma correta e apresentam oito passos para desenhar um modelo baseado em redes neurais, com o intuito de efetuar previsões de séries temporais. Zhang et al. (1998) fazem uma exposição das múltiplas opções a serem consideradas aquando da construção de uma RNA, sendo que um mau conjunto de opções pode levar a problemas de generalização do modelo fora da amostra. Estes autores, contudo, atribuem grande parte deste problema à metodologia de “tentativa e erro”, utilizada para a seleção da RNA na maioria dos estudos, que dizem levar também ao aparecimento de resultados contraditórios na literatura.

O primeiro problema que resulta da abordagem de tentativa e erro é a seleção das variáveis de entrada relevantes. Quando o número de variáveis de entrada é elevado, quando não existe conhecimento das relações lineares e não lineares entre as mesmas, ou quando se desconhece aquelas que têm maior capacidade de contribuir para a qualidade do modelo, pode tornar-se difícil encontrar as combinações de variáveis que maximizam a capacidade de previsão da RNA. Para resolver este problema, Dunis e Williams (2002) optam pela utilização de um algoritmo *stepwise* para a frente. Pacelli (2011), por sua vez, emprega o coeficiente de correlação de Pearson para eliminar variáveis e evitar o problema da multicolineariedade. Khashei e Bijari (2011b) utilizam “algoritmos de corte” para eliminar os pesos menos significativos da RNA e assim as variáveis a estes associados; no entanto, não explicitam o conteúdo dos algoritmos utilizados.

May et al. (2011) discutem também este problema, apresentando diversos algoritmos para a seleção de variáveis. No que respeita aos algoritmos de *stepwise* para a frente, os autores referem que estes podem ignorar relações entre as variáveis de entrada, altamente informativas conjuntamente, mas só marginalmente relevantes individualmente, fazendo com que o algoritmo possa ficar preso num mínimo local da função de erro. Algumas alternativas sugeridas passam pelo uso de coeficientes de correlação não lineares, que contornam os problemas oriundos da utilização de coeficientes lineares em problemas não lineares, ou por algoritmos genéticos que procuraram de forma eficiente as melhores combinações das variáveis de entrada, que reduzem segundo McNelis (2005) a probabilidade de a RNA ficar presa num mínimo local.

Trabalhos com redes neuronais, que pretendem analisar as relações não lineares entre variáveis de entrada e de saída, utilizam geralmente o coeficiente da informação mútua (May et al. 2011). O coeficiente da correlação da distância é uma alternativa recente. Clark (2013) conclui que ambos são capazes de encontrar relações não lineares entre variáveis, embora nenhum seja perfeito. Por sua vez Simon e Tibshirani (2011) demonstram que o coeficiente da correlação da distância apresenta resultados superiores ao da informação mútua.

O segundo problema que advém da abordagem de tentativa e erro é a escolha do número de nós e camadas escondidas da RNA. Esta é uma característica chave neste tipo de modelos, uma vez que são as camadas escondidas e os nós que captam os padrões entre as variáveis de saída e as variáveis de entrada.

Zhang et al. (1998) referem que, na generalidade, redes neuronais com poucos nós escondidos generalizam melhor, no entanto, nós a menos impedem que estas aprendam padrões nos dados. Estes autores apresentam ainda várias regras de ouro para a seleção destes dois parâmetros da RNA. Estas regras de ouro, todavia, podem ser refutadas porque ignoram facilmente a complexidade de que a função pode necessitar para se ajustar ao problema. Pacelli (2011) aborda este problema através de um algoritmo genético para escolher o número de nós, o número de camadas escondidas e ainda as funções de ativação, que conferem as propriedades não lineares à RNA, para cada camada escondida, tendo atingido boa performance no que diz respeito à generalização da RNA para dados diários.

Os modelos híbridos apareceram devido à necessidade de ultrapassar os problemas inerentes aos modelos estatísticos e aos modelos com base em inteligência artificial, numa tentativa de aumentar a capacidade de previsão.

Vários estudos sugerem modelos de hibridação. Zhang (2003b) utiliza um modelo híbrido univariado entre um modelo ARIMA e uma RNA para prever três séries temporais diferentes, entre as quais a taxa de câmbio GBP/USD, tendo provado a superioridade do modelo híbrido face aos modelos individuais e de passeio aleatório.

No entanto, os estudos referidos para a hibridação de modelos apresentam problemas, um deles prende-se com o facto de serem univariados. Os estudos que optam por hibridação univariada dependem apenas de valores passados da variável dependente. Assim, provavelmente, a capacidade de previsão pode ser potenciada se se providenciar um maior conjunto de informação ao modelo. Por outro lado, iniciar o modelo híbrido pela modelação da RNA, em vez

do ARIMA, poderá produzir melhores resultados. A racionalidade do início da hibridação pelo modelo linear prende-se no facto de as RNA possuírem também a capacidade de modelar, em parte, linearidades. Esta característica que pode levar a sobre-ajustamento aos dados por parte da RNA, justificando o início da modelação pelo modelo linear (Zhang 2003b). Porém, iniciar a hibridação multivariada por uma RNA pode captar melhor as relações não lineares existentes entre as variáveis, nomeadamente com a taxa de câmbio, permitindo depois ao modelo linear modelar as relações lineares em sobra nos erros. O problema está então em controlar o sobre-ajustamento da RNA.

Um outro problema está no facto de se considerar de forma inequívoca que os resíduos provenientes quer dos modelos lineares, quer dos modelos não lineares, contêm só não linearidades ou padrões lineares, consoante o caso, pelo que é necessário testar esta hipótese.

O modelo de Khashei e Bijari (2011a) tem em conta os dois últimos problemas referidos. No entanto, para casos multivariados, com um elevado número de entradas, a utilização desta metodologia pode aumentar desnecessariamente a dimensionalidade da RNA a estimar, assim como introduzir variáveis irrelevantes na mesma.

A dimensionalidade de uma RNA é resultante do aumento do número de pesos sinápticos (parâmetros da RNA) e do número de variáveis de entrada; tal como o problema da redundância das variáveis, também pode originar instabilidade na determinação do valor dos parâmetros, o que afetará a qualidade da previsão. Vários métodos foram propostos para eliminar estes dois problemas. Um deles emprega algoritmos genéticos para escolher conjuntamente o número de nós e as variáveis a considerar no modelo, encontrando simultaneamente uma solução ótima, ou próxima do ótimo. Makvandi et al. (2005) descreve uma metodologia para realizar tal seleção simultânea. No entanto, não tem em consideração as dificuldades que podem surgir quando a amostra é pequena, pois utiliza uma medida de ajustamento para comparar os modelos que o algoritmo analisa internamente, seleccionando o modelo com melhor ajustamento, não penalizando a sua dimensão face à dimensão da amostra. Tal não garante a seleção do modelo mais parcimonioso, ou seja não previne o aumento da dimensionalidade e consequente sobre-ajustamento da RNA seleccionada. Este problema pode ser resolvido utilizando um critério de informação para comparar os modelos.

Assim, a questão a colocar não deve estar limitada a qual a melhor ordem de hibridação de modelos, mas sim incluir também o problema do potencial de sobre-ajustamento ou sob-

ajustamento dos mesmos aos dados. O iniciar da hibridação por uma RNA multivariada, como se irá propor neste estudo, pode trazer vantagens, devido à capacidade desta para captar relações possivelmente não lineares entre as entradas e a saída.

Um último problema detetado é o défice de performance e de robustez do algoritmo de retropropagação tradicional, utilizado para treinar as RNA na maioria dos estudos, comparativamente a outros algoritmos (Falát 2012). De facto, Kamruzzaman e Sarker (2003) demonstram a superioridade de algoritmos com regularização bayesiana e do gradiente conjugado escalado com retropropagação (GCE) face ao algoritmo tradicional de retropropagação, para a previsão de taxas de câmbio semanais.

3 Construção da base de dados

Este trabalho estuda as taxas de câmbio EUR/USD e USD/JPY mensais e semanais, através de uma abordagem híbrida multivariada. Foram recolhidos dados mensais para o período de janeiro de 1999 a agosto de 2013, num total de 176 observações mensais para o EUR/USD, e entre janeiro de 1993 e agosto 2013, num total de 248 observações para o USD/JPY. Também recolhemos dados semanais, para o período entre janeiro de 2000 e Novembro de 2013, num total de 726 observações. O Quadro 1 apresenta as variáveis recolhidas.

Neste estudo não são analisadas as variáveis em nível, mas sim o logaritmo do quociente entre dois valores consecutivos, que dá aproximadamente a taxa de crescimento (semanais ou mensais, conforme o caso):

$$tc_n = \ln \left(\frac{S_n}{S_{n-1}} \right) \quad (7)$$

A fórmula (7), além de tornar as variáveis estacionárias, o que se verificou através do teste ADF, faz também com que os resultados da RNA não dependam de unidades de medida. Além disso, e embora não estritamente necessário, a estacionaridade facilita o processo de aprendizagem da RNA e de pré-processamento das entradas

Tabela 1 – Variáveis inicialmente consideradas para entradas da RNA multivariada

	EUR	USA	JPY
Desemprego	[1] (Sts.m.i6.n.uneh.rtt000.4.000)	[2] (Unratensa)	[2] (Lrunttttjpm156s)
Índice preço consumidor	[1] (lcp.m.u2.n.000000.4.anr)	[2] (Cpiaucns)	[2] (Jpnpciallminmei)
Índice Produção Industrial	[1] (Sts.m.i6.w.prod.ns0020.4.000)	[2] (Indpro)	[2] (Jpnproindmisme)
Taxa de Juro 1 Mês ^{1,2}	[1] (Fm.m.u2.eur.rt.mm.euribor1md)	[2] (Bbusd1m)	[2] (Jpy1mtd156n)
Taxa de Juro 3 Meses ^{1,2}	[1] (Fm.m.u2.eur.rt.mm.euribor3md)	[3] (usd3m)	[2] (Jpy3mtd156n)
Taxa de Juro 6 Meses ^{1,2}	[1] (Fm.m.u2.eur.rt.mm.euribor6md)	[3] (Bbusd6m)	[2] (Jpy6mtd156n)
Taxa de Juro 12 Meses ^{1,2}	[1] (Fm.m.u2.eur.rt.mm.euribor1yd)	[3] (Bbusd12)	[2] (Jpy12mtd156n)
Confiança de Consumidores	[1] (Rtd.m.s0.s.y_csccl.f)	[2] (cscicp03usm665s)	[2] (Cscicp03jpm665s)
Balança de Bens EUR-USA	[5] ³		
Balança de Bens JP-USA	[5]		
Preço Petróleo ²	[2] (Dcoilwtico)		
Preço Ouro ²	[4] (Gold)		
SP500 ²	[2] (SP500)		
Jpy/Usd ²	[2] (dexjpus)		
Eur/Usd ²	[2] (dexuseu)		

[1]-Eurostat, [2] FRED, [3] *Datastream* Internacional, [4]-Banco Mundial, [5] US Census Bureau

Notas: Entre parênteses curvos encontram-se os códigos das variáveis nas respetivas bases de dados. Todas as variáveis foram recolhidas mensalmente. 1) Euribor para Zona Euro e Libor para USA, JPY. 2) Variáveis também recolhidas semanalmente – média semanal. 3) Série construída através do somatório do saldo da balança de bens dos USA com cada país da zona euro a 15.

De seguida, os dados são divididos em duas partes, uma que se chama “dentro da amostra” e outra “fora da amostra”. A parte “fora da amostra” (a que alocámos 24 observações) serve unicamente para avaliar o desempenho do modelo com dados desconhecidos. Os dados dentro da amostra são utilizados para criar o modelo, podendo ainda ser divididos em dados de treino e dados de validação. A validação será utilizada pela RNA e pelo ARIMA para testar a potencial generalização do modelo fora da amostra. Para a amostra de validação reservaram-se 24 observações mensais e 55 semanais. As restantes observações servem para treino. É de

referir que perdemos 12 observações da amostra devido aos desfasamentos das variáveis que utilizaremos.

Fizemos uma pré-seleção das variáveis de entrada (incluindo os 12 desfasamentos) a testar no resto do estudo, utilizando para tal o coeficiente de correlação da distância. Foram selecionadas, de entre as entradas possíveis para cada período t , 25 variáveis de entrada de acordo com a correlação das variáveis em $t-1$ com a taxa de câmbio em t . Desta forma todas as restantes variáveis de entrada, com baixa correlação com a variável de saída, foram descartadas.

4 Metodologia: Modelos e Algoritmos

4.1 Rede Neuronal Artificial

As RNA podem ser descritas como uma tentativa de se replicar computacionalmente o funcionamento do cérebro humano. Os cálculos nestas redes são realizados alimentando as n

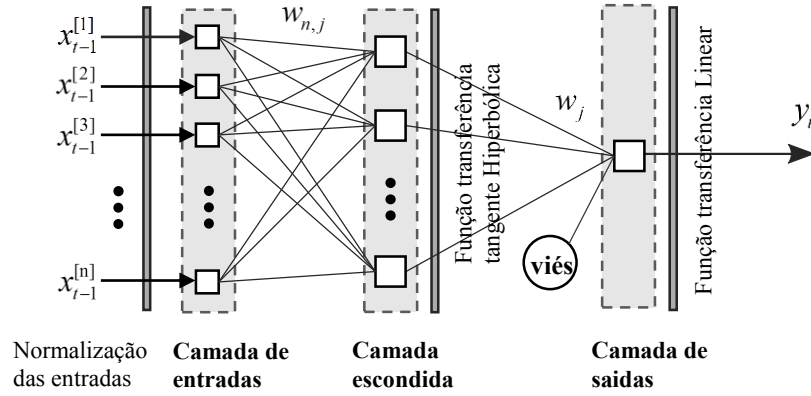


Figura 1. Estrutura genérica de uma Rede Neuronal Artificial utilizada neste estudo

entradas de uma camada de entradas, a partir da qual a RNA processa uma ou mais camadas escondidas, constituídas por q nós escondidos, e finalmente uma camada de saídas com m saídas, como demonstrado na Figura 1.

A Figura 1 é matematicamente representada pela equação:

$$y_t = \phi_0 \left(w_0 + \sum_{j=1}^q w_j \phi_1 \left(\sum_{i=1}^n w_{i,j} x_{t-1}^{[i]} \right) \right) + \varepsilon_t \quad (8)$$

onde w_0 é o viés da camada de saída e w_j ($j = 1, 2, \dots, q$) e $w_{i,j}$ ($i = 1, 2, \dots, n; j = 1, 2, \dots, q$) são os pesos sinápticos do modelo, responsáveis por armazenar a informação aprendida. Neste estudo, a função de ativação ϕ_1 utilizada nos nós escondidos, é a tangente hiperbólica – equação (9) abaixo –, a qual, por permitir uma normalização entre $[-1, 1]$, apresenta uma maior probabilidade de originar melhores resultados que outras funções (Karlik & Olgac 2010).

$$\phi_1(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \quad (9)$$

Desta forma, neste estudo, os dados a introduzir na RNA são normalizados entre $[-1, 1]$:

$$n(x) = \frac{2(x - x_{\min})}{x_{\max} - x_{\min}} - 1 \quad (10)$$

onde x é o valor a ser normalizado, x_{\max} e x_{\min} são respetivamente os valores máximo e mínimo de determinada série \mathbf{x} à qual pertence x .

A função ϕ_0 , utilizada no nó de saída, é uma função linear ($y = x$), tal como é normal neste género de estudos. $x_{t-1}^{[i]}$ pertence a um vetor de entradas \mathbf{x}_{t-1} , onde $t-1$ indica o período na amostra e i ($i = 1, 2, \dots, n$) indica a variável de entrada. Por fim y_t e ε_t são, respetivamente, o valor e o erro no período t . A equação (8) é assim uma função não linear que realiza o mapeamento entre observações passadas e futuras. Neste estudo, o viés apenas é introduzido na camada de saídas, dado que as vantagens da sua introdução, no que respeita à previsão de séries temporais, revelam-se inexistentes, ou haverá mesmo desvantagens (Zhang & Hu 1998). Desta forma, obtém-se também um modelo mais parcimonioso, seguindo o princípio da navalha de Occam.

Assim a equação (8) pode ser representada por:

$$y_t = f(\mathbf{x}_{t-1}, \mathbf{w}) + \varepsilon_t \quad (11)$$

onde \mathbf{w} é o vetor de pesos sinápticos e do viés e f a função não linear determinada pela RNA.

O treino da RNA (estimação de \mathbf{w}) requer a utilização de um algoritmo apropriado para o efeito. Apesar de existirem diversos algoritmos que podem ser utilizados nesta estimação, nenhum é capaz de garantir que se atinge o mínimo global da função f . Neste estudo utiliza-se o algoritmo GCE, implementado na função ‘*trainscg*’ do MATLAB. É um método robusto

para estimar \mathbf{w} , e a sua performance é particularmente boa em problemas com um grande número de pesos (Zakaria 2009), como é o problema em estudo.

O algoritmo utiliza geralmente a amostra de validação para testar o vetor \mathbf{w} treinado, procurando minimizar os erros na amostra de validação, de acordo com uma medida de ajustamento. A este processo chama-se “de paragem prematura”, e tem o intuito de evitar o sobre-ajustamento à amostra de treino. Para tal tarefa, neste estudo é utilizado o MSE:

$$MSE = \frac{1}{T} \sum_{i=1}^T (\hat{y}_i - y_i)^2 \quad (12)$$

onde T é o número de observações da amostra.

No entanto, a utilização deste método em amostras pequenas continua a poder levar a problemas de sobre-ajustamento. Um outro problema advém da própria utilização da amostra de validação, que remove as observações mais recentes da informação utilizada para realizar previsão.

Uma alternativa para aumentar a capacidade de generalização da RNA fora da amostra passa por treinar a RNA utilizando a técnica da “regularização”. A principal vantagem do uso da regularização na RNA é que esta não necessita de uma amostra de validação para treinar o modelo, sobrando assim mais observações para o seu treino. A regularização impõe o princípio da navalha de Occam ao treino da RNA, pois não se minimizam só os erros, mas sim uma combinação destes e de \mathbf{w} :

$$C = \alpha E_D + (1 - \alpha) E_w \quad (13)$$

onde E_D é o MSE, E_w é a soma do quadrado dos pesos sinápticos e do viés, e $\alpha \in [0,1]$ é o parâmetro de regularização. Este processo é também conhecido por “decaimento dos pesos”. O problema desta abordagem está no facto de se desconhecer qual o melhor valor do parâmetro α a utilizar para treinar dentro da amostra.

Devido facto de o uso do algoritmo GCE não garantir que se encontrará o mínimo global, na prática utiliza-se este algoritmo várias vezes no mesmo problema, escolhendo (possivelmente de forma aleatória) um valor inicial para \mathbf{w} diferente em cada uma dessas vezes. Este estudo, implementando código para tal, inicializa \mathbf{w} 30 vezes e combina os 5 melhores vetores dos resultados das saídas das RNA, através de uma média simples.

A discussão da construção da RNA e posterior estimação de α é feita no ponto 4.2.

4.2 Metodologia para construção de uma RNA

No nosso estudo das taxas de câmbio EUR/USD e USD/JPY, utilizámos um algoritmo genético para encontrar a melhor combinação de variáveis de entrada, número de nós e camadas escondidas para o tipo de RNA descrita em 4.1. Depois de determinada, desse modo, a estrutura, estimou-se o parâmetro α , a utilizar no treino com regularização da RNA.

4.2.1 Algoritmo genético adaptativo proposto

Os algoritmos genéticos (AG) são uma técnica de procura adaptativa utilizada em problemas de otimização. Há um paralelismo entre os AG e a teoria da seleção natural de Darwin, que se baseia na ideia de que os indivíduos mais fortes têm uma maior probabilidade de sobrevivência.

Os AG funcionam com um conjunto de vetores candidatos a serem o ótimo e ao qual se chama “população”. A população é criada aleatoriamente no início do algoritmo, sendo constituída por um número N (par) de “indivíduos”. Cada indivíduo corresponde a um vetor com r genes (coeficientes) aleatórios:

$$\begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_r \end{pmatrix}_1 \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_r \end{pmatrix}_2 \dots \begin{pmatrix} \Phi_1 \\ \Phi_2 \\ \vdots \\ \Phi_r \end{pmatrix}_N \quad (14)$$

Em termos gerais, o algoritmo consiste em repetir os seguintes passos um certo número de vezes (“gerações”):

- 1) Cada indivíduo é avaliado com base numa função que reflete o desempenho da RNA caso se utilizem os coeficientes associados a esse indivíduo;
- 2) Os indivíduos com melhor avaliação são selecionados como “sobreviventes” para a próxima iteração;
- 3) Formam-se novos indivíduos resultantes da mistura de coeficientes de indivíduos diferentes (cruzamento), ou de alterações aleatórias (mutações) dos coeficientes dos indivíduos existentes;
- 4) O algoritmo volta ao primeiro passo, tendo como população os sobreviventes e os novos indivíduos resultantes de cruzamentos e de mutações.

O ótimo será o indivíduo da última geração que possuir a melhor avaliação.

Foi necessário escrever um código para implementarmos no MATLAB este algoritmo, fazendo corresponder a cada membro da população uma estrutura diferente para a RNA em função dos genes – número binário, 0 ou 1 (bit) – desse indivíduo.

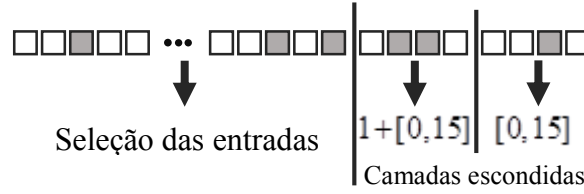


Figura 2. Estrutura de um indivíduo no algoritmo genético proposto

A Figura 2 representa a estrutura genética de um indivíduo. Os últimos 8 genes são responsáveis pela seleção do número de camadas e nós escondidos na RNA. Recorde-se que 4 bits na base dois podem representar números decimais entre 0 e 15. Garante-se que a primeira camada escondida possui sempre pelo menos um nó (e um máximo de 16) adicionando 1 a qualquer valor decimal resultante. Os restantes genes são responsáveis pela seleção das entradas na RNA, indicando se uma variável é ou não introduzida no vetor de entradas da RNA. Zhang et al. (1998) referem que a utilização de duas camadas escondidas pode ser vantajosa; no entanto, tal prática não é adotada na maioria dos estudos com séries temporais. No nosso caso, permite-se ao AG selecionar ou descartar a segunda camada escondida.

Como dissemos, a informação contida nos genes de cada indivíduo corresponde uma RNA. A capacidade de generalização da RNA é avaliada na amostra de validação através do método da paragem prematura descrito em 4.1.

Os indivíduos são depois comparados, dentro da amostra, através do critério de Hannan-Quinn corrigido para amostras pequenas, na tentativa de obter o modelo mais parcimonioso:

$$HQc = T \ln \left(\frac{\sum_{i=1}^T (\hat{y}_i - y_i)^2}{T} \right) + \frac{2K \ln(\ln(T))}{T - K - 1} \quad (15)$$

Nesta equação, y_i é o valor da taxa de câmbio (transformada) na data i , \hat{y}_i é o valor estimado pela RNA, T é o número de observações e K é o número de parâmetros. Neste contexto, uma amostra é considerada pequena quando a “regra de ouro” $T / K < 40$ é verdadeira, onde

T é o tamanho da amostra e $K = n \times q + q + 1$, ou seja o número de pesos sinápticos mais o viés.

O próximo passo é a seleção da próxima geração de indivíduos. É garantida a passagem dos dois melhores indivíduos à geração seguinte (elitismo), sendo que todos os indivíduos incorrem depois em processos de cruzamento e mutação, como representado na Figura 3 e explicado de seguida. Dos N membros da população, 70% são selecionados aleatoriamente, dois a dois, para originarem um terceiro membro da população através de cruzamento, sendo que os indivíduos que possuem menor valor de HQc , apresentam maior probabilidade de serem selecionados. Os restantes sofrem mutação em cada um dos seus genes com probabilidade de Pm , definido da seguinte forma:

$$\begin{cases} Pm = 0.5, & \text{se } f' > f_{med} \\ Pm = 0.03 \frac{|f'| - |f_{min}|}{|f_{med}| - |f_{min}|}, & \text{se } f' \leq f_{med} \end{cases} \quad (16)$$

onde f' é o valor de HQc do indivíduo que irá sofrer mutação, f_{med} é o valor médio de HQc dos indivíduos que sofrem mutação, e f_{min} é o valor de HQc do melhor indivíduo na população. Assim indivíduos próximos do melhor sofrem pequenos ajustes, na tentativa de os melhorar, enquanto os piores indivíduos sofrem mutações em aproximadamente 50% dos seus genes. Esta característica adaptativa permite ao algoritmo encontrar, na maioria dos casos, soluções próximas do ótimo.

Os novos indivíduos, resultantes dos processos de elitismo, cruzamento e mutação, substituem os piores dentro da população e o AG volta ao início com esta nova geração. Quando o número de gerações atingir o limite imposto – neste estudo, 100 gerações, ou, em alternativa, existirem 20 gerações seguidas em que o melhor indivíduo se mantém o mesmo –, o algoritmo para e guarda a estrutura da RNA, dada pelo melhor indivíduo, a utilizar na previsão.



Figura 3. Representação dos processos de cruzamento e mutação utilizados pelo algoritmo genético

Na prática não se pode garantir que a estrutura selecionada pelo AG é a melhor, uma vez que este não percorre todas as possibilidades. Para aumentar a probabilidade de o resultado estar próximo do ótimo, aplica-se o AG cinco vezes, com populações iniciais diferentes, e seleciona-se a estrutura, entre as cinco RNA resultantes, que minimiza o critério de Hannan-Quinn corrigido para amostras pequenas. Não foi utilizado um número maior de inicializações do AG devido ao tempo que este requer para processar a estrutura da RNA, que com dois *cores* a 2.13GHz demora em média 5 horas para cada inicialização.

4.2.2 Regularização no treino da RNA

Depois de encontrado o potencial melhor modelo através do algoritmo genético, procede-se à estimação do parâmetro α da equação (13), necessário para treinar a RNA. Tal procedimento é feito estimando a estrutura da RNA encontrada para diferentes valores de α , e realizando a validação cruzada dentro da amostra, que é uma abordagem frequentemente utilizada para este fim (Plasmans et al. 1998). Nesta técnica, o treino e a validação da RNA processam-se de forma iterativa e sequencial, em múltiplas subamostras, minimizando a média do MSE dessas subamostras.

Contudo, para séries temporais, tem de se ter o cuidado adicional de manter a estrutura temporal dos dados. Ou seja, a RNA é treinada, inicialmente, numa amostra com k observações e a validação é realizada, neste estudo, através do MSE nas observações $k+1$, $k+2$ e $k+3$, sendo o valor do MSE guardado. De seguida o processo treina a mesma estrutura da RNA com $k+3$ observações e valida em $k+4$ a $k+6$. O processo repete-se até não se possuírem mais dados dentro da amostra. O valor de α que minimizar a média do MSE é utilizando posteriormente para realizar previsões fora da amostra. O processo descrito aumenta a probabilidade de que o indivíduo selecionado no final do algoritmo genético seja aquele que cria a RNA com melhor desempenho fora da amostra.

4.3 Modelos de referência

Vários modelos de referência são utilizados neste estudo para comparar o modelo híbrido proposto no ponto 4.5 a outras abordagens na literatura, nomeadamente abordagens univariadas.

4.3.1 Modelo ARIMA

Introduzido por Box e Jenkins, o modelo ARIMA tem sido uma das abordagens de maior popularidade para a realização de previsões. Neste modelo univariado, o valor futuro da variável é assumido como sendo uma função linear de observações passadas e de erros:

$$y_t = \theta_0 + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \dots - \theta_q \varepsilon_{t-q} \quad (17)$$

onde o valor de y_t e e_t são respetivamente o valor observado e o erro no período t . φ e θ são parâmetros do modelo. p e q são números inteiros, frequentemente referidos como ordens autorregressivas e de média móvel, respetivamente.

Neste estudo, a metodologia utilizada para selecionar os modelos ARIMA procurou selecionar o modelo mais parcimonioso, ou seja, o modelo mais simples que não apresenta piores resultados dentro da amostra comparado a outros. Para tal, numa primeira fase de pré-seleção, usou-se o HQc dado pela equação (15), sendo neste caso $K=p+q$. Dos modelos pré-selecionados, o que minimiza o MSE na amostra de validação foi o selecionado.

A estacionariedade da série é uma condição necessária na construção de um modelo ARMA, caracterizada por uma média e estrutura de autocorrelação constantes ao longo do tempo. Para a maioria das séries temporais tal condição é satisfeita depois da transformação da série em taxas de crescimento.

4.3.2 Modelo híbrido convencional para séries temporais univariadas

Na literatura, o modelo híbrido mais utilizado combina um modelo ARIMA e uma RNA, de forma univariada. Nestes trabalhos, os autores aplicam o modelo ARIMA inicialmente para modelar os dados iniciais e de seguida modelam os resíduos obtidos, e considerados não line-

ares, através de uma RNA. Neste estudo irá utilizar-se o modelo proposto por Khashei e Bijari (2011a), programámos em MATLAB. Este tipo de modelo híbrido pode ser ilustrado como:

$$y_t = L_t + N_t \quad (18)$$

onde L_t e N_t são, respetivamente, as componentes lineares e não lineares do modelo híbrido univariado. Numa primeira fase calcula-se a previsão para t do modelo ARIMA, de tal forma que:

$$L_t = \hat{L}_t + \varepsilon_t \quad (19)$$

onde \hat{L}_t é o valor da previsão e ε_t o erro para o período t . Numa segunda fase, uma RNA é utilizada para modelar a variável de interesse, y_t . A informação usada inclui os desfasamentos dos resíduos do modelo ARIMA e a previsão produzida por esse modelo, assim como desfasamentos da variável de entrada inicial, de tal forma que:

$$y_t = f(\varepsilon_{t-1}, \dots, \varepsilon_{t-n}, \hat{L}_t, z_{t-1}, \dots, z_{t-m}, \mathbf{w}) \quad (20)$$

onde f é a função não linear determinada pela RNA, e z_{t-1} é o primeiro desfasamento da variável de entrada.

4.3.3 Modelo de Passeio Aleatório

O modelo de passeio aleatório é utilizado neste estudo como modelo de comparação da performance para os restantes. Este é definido por:

$$y_{t+1} = y_t + \varepsilon_{t+1} \quad (21)$$

onde y_t é o valor no período t . A previsão para o período seguinte será o valor corrente. Este modelo é um subcaso do modelo ARIMA referido em 4.3.1., mais concretamente, o ARIMA(0,1,0).

4.4 Testar a não linearidade: Testes BDS e LWG

O teste de Brock, Dechert e Scheinkman (BDS) e o teste de Lee, White e Granger (LWG) são utilizados neste estudo para examinar a existência de não linearidades ou linearidades não modeladas nos resíduos dos modelos, e avaliar a possibilidade de hibridação. Ao que se pôde apurar, estes dois testes nunca foram utilizados em conjunto para avaliar a possibilidade de

hibridação entre dois modelos, uma vez que geralmente os estudos assumem que uma série temporal possui padrões lineares e não lineares. Contudo, esse não é obrigatoriamente o caso, como se comprovará neste estudo.

O teste BDS testa, para uma dimensão escolhida, a hipótese nula de determinada série ser independente e identicamente distribuída (i.i.d.), contra uma hipótese alternativa não definida. No entanto, é capaz de indicar a existência de linearidades, não linearidades ou série caótica no caso de rejeição da hipótese nula (Matilla-García & Ruiz 2005), embora não seja possível especificar qual dos casos ocorre.

Por sua vez, o teste LWG testa a existência de não linearidades não modeladas. Não distingue, contudo, se, no caso de rejeição da hipótese nula, a série é i.i.d. ou se possui padrões lineares.

Aplicando o teste BDS aos resíduos de determinado modelo é possível testar se ainda existem padrões não modelados nos mesmos, quer sejam lineares ou não. A posterior aplicação do teste LWG permite tirar conclusões em relação ao tipo de padrões não modelados.

Esta combinação de testes é importante, pois, no caso de os resíduos de determinado modelo serem i.i.d., não é plausível modelar os mesmos através de um modelo linear ou não linear.

4.5 O modelo híbrido multivariado proposto

O modelo híbrido multivariado que se propõe neste estudo, segue uma abordagem para hibridação inversa aos tradicionais modelos de hibridação univariados, cujas razões para tal já foram discutidas na secção 2. Assim, o modelo proposto pode ser representado por:

$$y_t = N_t + L_t \quad (22)$$

onde L_t e N_t são, respetivamente, as componentes lineares e não lineares do modelo híbrido multivariado. Primeiramente, estima-se a previsão para t da RNA multivariada, de tal forma que:

$$y_t = \hat{N}_t + \varepsilon_t \Leftrightarrow \varepsilon_t = y_t - \hat{N}_t \quad (23)$$

onde \hat{N}_t é o valor da previsão e ε_t o erro para o período t . De seguida, verifica-se a inexistência de padrões não lineares e a existência de padrões lineares em ε através do teste descrito em 4.4. Caso tal se confirme, ε é modelado através de um modelo ARIMA:

$$\varepsilon_t = f(\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-n}) + \mu_t \quad (24)$$

onde f é a função linear determinada pelo modelo ARIMA e μ_t o erro que deve ser i.i.d. Por fim, a hibridação é realizada através de uma relação aditiva entre as componentes lineares e não lineares da série:

$$\hat{y}_t = \hat{N}_t + \hat{L}_t \quad (25)$$

onde \hat{L}_t corresponde ao valor da previsão para ε_t dada por (24). Considerar uma relação aditiva, ao invés de uma relação não linear, entre as duas componentes da série para realizar a hibridação dos dois modelos é um procedimento criticado por Khashei e Bijari (2011a). No entanto, testou-se a utilização de uma RNA para realizar a hibridação entre as duas componentes e obtiveram-se resultados comparáveis aos obtidos com a relação aditiva, pelo que, pelo menos para as séries consideradas, a existência de uma relação aditiva entre os modelos não conduz a resultados inferiores.

5 Resultados empíricos

Nesta secção investiga-se a capacidade de previsão do modelo proposto face aos restantes modelos utilizados no estudo. Para comparar a performance dos modelos é necessário estimá-los em dados desconhecidos para eles, ou seja fora da amostra usada na sua construção. Assim, todos os modelos foram comparados num período idêntico fora da amostra, de forma a se poderem realizar inferências relativamente à capacidade preditiva dos mesmos.

As previsões são realizadas um período à frente, estimando o modelo para as primeiras n_0 observações. De seguida, os parâmetros do modelo são estimados com n_0+1 observações e a próxima previsão um período à frente é calculada. O processo repete-se enquanto $n_0+k < T$, onde T é o tamanho da série considerada e k a iteração.

São consideradas três medidas para a comparação dos modelos: o RMSE, o SMAPE e o CDC (ver Apêndice A).

Devido à flexibilidade das RNA, já discutida anteriormente, comparar, dentro da amostra, os erros de previsão da RNA e os erros de previsão doutros tipos de modelos tem pouco interesse, devido à facilidade que as RNA têm em apresentar performance superior a outros modelos dentro da amostra. Por isso, este estudo foca-se nos resultados obtidos fora da amostra.

Tabela 2 - Estatísticas para a previsão fora da amostra da taxa de câmbio EUR/USD Mensal

	ARIMA(0,1,0)	RNA [6-9-8-1]	RNA[6-9-8-1] + ARIMA (5,0,6)	ARIMA(1,1,7)	ARIMA(1,1,7)+ RNA [5,6,4,1]
RMSE	0.0379	0.0392	0.0374	0.0388	0.0387
SMAPE	0.0113	0.0115	0.0109	0.0107	0.0108
CDC	38%	42%	46%	42%	42%

Nota: RNA+ARIMA, representa o modelo híbrido multivariado proposto neste estudo, enquanto que ARIMA+RNA corresponde ao modelo híbrido univariado de Khashei e Bijari. Entre parênteses retos encontra-se a estrutura da RNA [a,b,c,d], onde 'a' corresponde ao número de entradas, 'b' e 'c' ao número de nós na primeira e segunda camadas escondidas (respetivamente) e 'd' ao número de saídas da RNA. Entre parênteses curvos (e, f, h) encontram-se a ordem auto regressiva, de integração e de média móvel, respetivamente, do modelo ARIMA.

Para os dados mensais da taxa de câmbio EUR/USD, a Tabela 2 mostra que o modelo proposto apresenta melhores resultados quando se utiliza o RMSE ou o CDC como critério de comparação, enquanto o modelo univariado ARIMA(1,1,7) apresenta o melhor resultado quando se utiliza o SMAPE. De forma geral, todos os modelos parecem apresentar resultados superiores ao passeio aleatório no que diz respeito à capacidade de prever a direção do movimento da série (CDC). No entanto, o modelo proposto é o único que apresenta melhores resultados que o passeio aleatório em todos os três testes considerados.

A RNA multivariada é, nos testes considerados, excetuando o CDC, o pior modelo testado, o que pode significar a não existência de relações não lineares entre as entradas e a taxa de câmbio EUR/USD. A aplicação do teste descrito em 4.4 aos resíduos da RNA multivariada demonstra que não existem não linearidades por modelar e que os resíduos são passíveis de serem modelados através de um modelo linear. Ao realizar a hibridação desta RNA com um modelo ARIMA, verifica-se um aumento da capacidade preditiva, em todos os três testes considerados, relativamente à RNA multivariada.

De facto, aplicando também o teste descrito em 4.4 aos resíduos do modelo univariado ARIMA, conclui-se que estes não possuem não linearidades por modelar. Apesar do resultado deste teste, foi realizada a hibridação deste modelo ARIMA com uma RNA, seguindo o mode-

lo de Kashei e Bijari, para se verificar que este não se traduz em melhorias substanciais na capacidade preditiva relativamente ao ARIMA (1,1,7). Verifica-se mesmo para o SMAPE um valor mais elevado. Assim, nesta série e para o período testado, é de concluir que com dados mensais o par EUR/USD não apresenta não linearidades para modelar.

Tabela 3 - Estatísticas para a previsão fora da amostra da taxa de câmbio JPY/USD Mensal

	ARIMA(0,1,0)	RNA [5-11-6-1]	RNA[5-11-6-1]+ ARIMA (2,0,5)	ARIMA(2,1,3)	ARIMA(2,1,3)+ RNA [10,6,9,1]
RMSE	2.285	2.368	2.176	2.172	2.213
SMAPE	0.0106	0.0108	0.0101	0.099	0.0104
CDC	62%	62%	62%	58%	58%

Nota: ver nota da tabela 2.

A Tabela 3 apresenta as estatísticas relativas ao par JPY/USD mensal. O modelo híbrido proposto revela nesta série resultados semelhantes, embora com pior ajustamento, relativamente ao modelo ARIMA(2,1,3). Este último e o modelo híbrido univariado exibem, no entanto, menor capacidade de prever a direção do movimento da série, relativamente aos restantes.

É de notar a redução da qualidade do ajustamento do modelo híbrido univariado de Kashei e Bijari face ao modelo ARIMA univariado. De facto, da aplicação do teste descrito em 4.4, verifica-se que os resíduos do modelo ARIMA(2,1,3) são i.i.d., pelo que a modelação destes através de um modelo não linear não é a adequada.

Comparativamente à taxa de câmbio EUR/USD, a RNA multivariada apresenta também neste par de moedas o pior resultado de entre todos os modelos para as medidas de ajustamento RMSE e SMAPE. Mais uma vez, tal resultado pode dever-se à inexistência de não linearidades para modelar. De facto, quando realizada a hibridação através do modelo híbrido multivariado proposto, os resultados podem ser equiparados, como já referido, ao modelo linear ARIMA(2,1,3), o que vez pode ser evidência da não existência de não linearidades para modelar nesta série mensal. Dado, o que confere poder preditivo ao modelo híbrido, face ao passeio aleatório, é a modelação dos erros da RNA multivariada através de um ARIMA(2,0,5).

Tabela 4 - Estatísticas para a previsão fora da amostra da taxa de câmbio EUR/USD Semanal

	ARIMA(0,1,0)	RNA [8,14,7,1]	RNA[8,14,7,1]+ ARIMA (2,0,0)	ARIMA(2,1,5)	ARIMA(2,1,5)+ RNA [8,10,5,1]
RMSE	0.01170	0.01148	0.01131	0.01144	0.01144
SMAPE	0.0036	0.0035	0.0034	0.0036	0.0036
CDC	54%	54%	54%	54%	50%

Nota: ver nota da tabela 2.

Na Tabela 4, verifica-se, para dados semanais da taxa de câmbio EUR/USD, que o modelo híbrido multivariado proposto apresenta melhor ajustamento que os restantes, apesar de nenhum modelo se destacar em termos de capacidade de previsão da direção da série.

Tal como com dados mensais para a taxa de câmbio EUR/USD semanal, não parece existir um modelo que se destaque face ao de passeio aleatório. A aplicação de um modelo híbrido univariado a esta série semanal não se traduz em aumento do poder preditivo relativamente ao ARIMA(2,1,5). O teste BDS aos resíduos deste último leva à rejeição da hipótese nula do mesmo. No entanto, o teste LWG rejeita a hipótese da existência de não linearidades nos erros do modelo ARIMA, pelo que estes serão provavelmente caóticos, ou existe a possibilidade de ainda existirem linearidades por modelar. Assim, aplicou-se também o teste de Ljung-Box aos resíduos deste ARIMA, o qual não rejeita a hipótese nula destes serem i.i.d., pelo que se comprova uma adequada especificação do modelo ARIMA, cujos resíduos não são passíveis de serem modelados por um modelo não linear. Verificou-se mesmo a redução da capacidade de previsão da direção da série aquando da hibridação univariada.

Tabela 5 - Estatísticas para a previsão fora da amostra da taxa de câmbio JPY/USD Semanal

	ARIMA(0,1,0)	RNA [7,16,5,1]	RNA[7,16,5,1]+ ARIMA(3,0,2)	ARIMA(0,1,0)	ARIMA(0,1,0)+ RNA [5,7,4,1]
RMSE	1.0833	0.8401	0.7970	1.0833	0.9102
SMAPE	0.0041	0.0040	0.0031	0.0041	0.0040
CDC	67%	75%	75%	67%	75%

Nota: ver nota da tabela 2.

A Tabela 5 apresenta estatísticas referentes ao par de moedas JPY/USD semanal. De todos os pares estudados, este é aquele que apresenta inequivocamente não linearidades nas relações entre as variáveis. O melhor modelo ARIMA selecionado para esta taxa de câmbio corresponde ao passeio aleatório. No entanto, a hibridação deste com uma RNA, segundo o modelo de Kashei e Bijari, prova a existência de não linearidades nos resíduos do modelo

ARIMA selecionado, verificando-se um aumento da capacidade preditiva do modelo híbrido face ao ARIMA.

O modelo híbrido multivariado proposto é o melhor em todos os três testes considerados. É de notar que, por si só, a RNA multivariada apresenta um ajustamento superior ao do modelo híbrido univariado e ao do passeio aleatório, o que sugere a existência de relações não lineares entre as entradas da RNA e a taxa de câmbio.

Por fim, a aplicação do teste descrito e 4.4 a todos os modelos híbridos modelados nesta secção demonstra que os seus resíduos não apresentam padrões por modelar.

6 Conclusão

Este estudo avaliou o uso de um novo modelo híbrido multivariado entre uma RNA e um modelo ARIMA para a previsão das taxas de câmbio EUR/USD e JPY/USD. A performance foi comparada com a de um modelo de passeio aleatório e a de um modelo híbrido univariado.

Excetuando a taxa de câmbio JPY/USD semanal, as previsões dadas pelo passeio aleatório não se revelaram muito piores que as do melhor modelo alternativo. De facto, estes resultados estão de acordo com o disposto na bibliografia analisada no ponto 2, nomeadamente no que diz respeito a existirem indícios de a taxa de câmbio EUR/USD ser passeio aleatório e à incapacidade das RNA para a modelação de dados mensais. Esta incapacidade das RNA parece, no entanto, dever-se à não existência de padrões não lineares nestas séries, quer quando analisadas com modelos multivariados, quer com modelos univariados. Conclui-se assim que, para as séries mensais consideradas, não existem padrões não lineares passíveis de serem modelados, pelo que a hipótese de Zhang (2003b), de que as séries são constituídas por uma componente linear e outra não linear, não se verifica.

Devido à complexidade dos modelos híbridos, para taxas de câmbio mensais, não parece ser aconselhada a utilização destes modelos. Assim, a opção por um modelo mais simples deve ser preferida em detrimento dos modelos híbridos. De facto, deve sempre ser selecionado, como melhor modelo, aquele que sendo mais simples não se traduz em piores resultados do que outros mais complexos.

Para a taxa de câmbio JPY/USD semanal, onde se encontraram indícios de padrões não lineares, o modelo híbrido multivariado proposto destacou-se face aos restantes. Assim, a

existirem padrões não lineares nas séries, as previsões do modelo proposto parecem apresentar melhor ajustamento. Este resultado permite ainda inferir sobre a potencial vantagem da utilização de um maior conjunto de variáveis na previsão de taxas de câmbio com horizonte temporal semanal ou mesmo diário.

No entanto, quando estes padrões não lineares não existem, modelar um modelo híbrido complexo não parece trazer vantagens. Desta forma, é recomendável testar a existência destes padrões nas séries antes de se proceder à modelação de modelos não lineares e posterior hibridação com modelo lineares, ou *vice versa*. Lixia (2009) apresenta uma metodologia para tal. No entanto, para as séries consideradas, a combinação de testes proposta em 4.4 aos erros de um modelo ARIMA parece permitir, pelo menos numa análise univariada, inferir sobre a vantagem da utilização de uma abordagem não linear ou híbrida. No entanto, para podermos generalizar esta conclusão, necessitaríamos de realizar mais testes com séries diferentes.

Há ainda outras hipóteses que podem ser exploradas, por exemplo, considerando indicadores técnicos, juntamente com dados macroeconómicos, como entradas no modelo híbrido multivariado, de forma a diversificar a informação disponível para o modelo e potenciar a qualidade da previsão. É necessário ainda generalizar os resultados obtidos para o modelo proposto a outras séries, nomeadamente semanais e diárias. Melhorias da metodologia de construção da RNA podem também ser obtidas aplicando o algoritmo de regularização bayesiana ao treino da RNA, que possui o potencial de apresentar resultados equivalentes ou melhores com a vantagem de não depender da correta especificação da estrutura da RNA.

No geral, ignorando a complexidade do modelo e da metodologia proposta, assim como o tempo necessário à sua estimação (dependente da capacidade de processamento do computador médio atual), o modelo proposto apresenta melhorias face aos restantes, quando se verifica a existência de não linearidades para modelar.

7 Bibliografia

- Cheung, Y.-W., Chinn, M.D. & Pascual, A.G., 2005. Empirical exchange rate models of the nineties: Are any fit to survive? *Journal of International Money and Finance*, 24(7), pp.1150–1175.
- Clark, M., 2013. *A comparison of correlation measures*, Center for Social Research university for notre dame.

- Dunis, C.L. & Williams, M., 2002. Modelling and trading the EUR / USD exchange rate : Do neural network models perform better ? *Center for International Banking, Economics and Finance*, pp.1–24.
- Falát, L., 2012. Numerical analysis of learning algorithms and strategies of multilayer perceptron neural network. *International Virtual Conference*, pp.12–17.
- Gradojevic, N. & Yang, J., 2000. The application of artificial neural networks to exchange rate forecasting: The role of market microstructure variables. *Financial Markets Department, Bank of Canada*, p.26.
- Hann, T.H. & Steurer, E., 1996. Much ado about nothing? Exchange rate forecasting: neural networks vs. linear models using monthly and weekly data. *Neurocomputing*, pp.323–339.
- Huang, W. et al., 2006. Comparisons of the different frequencies of input data for neural networks in foreign exchange. In V. N. Alexandrov et al., eds. *Computational Science – ICCS*. pp. 517–524.
- Jamal, A.M.M. & Sundar, C., 1998. Modeling exchange rates with neural networks. *Journal of Applied Business Research*, 14(1).
- Kaasra, I. & Boyd, M., 1996. Design ANN for forecast economic time series. *Neurocomputing*, (10), pp.215–236.
- Kamruzzaman, J. & Sarker, R. a., 2003. Forecasting of currency exchange rates using ANN: a case study. *International Conference on Neural Networks and Signal Processing*, pp.793–797 Vol.1.
- Karlik, B. & Olgac, A.V., 2010. Performance analysis of various activation functions in generalized MLP architectures of neural networks. *International Journal of Artificial Intelligence And Expert Systems*, 1(4), pp.111–122.
- Khashei, M. & Bijari, M., 2011a. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. *Applied Soft Computing*, 11(2), pp.2664–2675.
- Khashei, M. & Bijari, M., 2011b. Exchange rate forecasting netter with hybrid artificial neural networks models. *Journal of Mathematical and Computational Science*, (1), pp.103–125.
- Koloch, G., Rubaszek, M. & Skrzypczyński, P., 2011. Forecasting the polish zloty with non-linear models. *NATIONAL BANK OF POLAND*, (81).
- Lixia, L., 2009. Testing for nonlinearity and chaoticity in exchange rate time series. In H. Zhang, ed. *Marketing Science Innovations and Economic Development: Proceedings of 2009 Summit International Marketing Science and Management Technology Conference*. pp. 242–245.
- Makvandi, P., Jassbi, J. & Khanmohammadi, S., 2005. Application of genetic algorithm and neural network in forecasting with good data. In *Proceedings of the 6th World Scientific and Engineering Academy and Society (WSEAS) international conference on neural networks*. pp. 56–61.

- Matilla-García, M. & Ruiz, J.R., 2005. Aplicabilidad del test BDS al análisis de series económicas. *Estudios de Economía Aplicada*, 23(23), pp.507–519.
- May, R., Dandy, G. & Maier, H., 2011. Review of input variable selection methods for artificial neural networks. In K. Suzuki, ed. *Artificial Neural Networks - Methodological Advances and Biomedical Applications*. InTech, p. 362.
- McNelis, P.D., 2005. *Neural networks in finance: Gaining predictive edge in the market*, Elsevier Academic Press.
- Meese, R.A. & Rogoff, K., 1983. Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, (14), p.21.
- Pacelli, V., 2011. An artificial neural network model to forecast exchange rates. *Journal of Intelligent Learning Systems and Applications*, 03(02), pp.57–69.
- Panda, C. & Narasimhan, V., 2007. Forecasting exchange rate better with artificial neural network. *Journal of Policy Modeling*, 29(2), pp.227–236.
- Plasmans, J., Verkooijen, W. & Daniels, H., 1998. Estimating structural exchange rate models by artificial neural networks. *Applied Financial Economics*, (8), pp.541–551.
- Seliem, M.L., 2006. *Foreign Exchange forecasting using artificial neural network as data mining too*.
- Simon, N. & Tibshirani, R., 2011. *Comment on “Detecting novel associations in large data sets” by Reshef et al.*,
- Zakaria, Z., 2009. A study on neural network training algorithm for multiface detection in static images. *Psp’s Research Digest*, (1), pp.12–16.
- Zhang, G. & Hu, M.Y., 1998. Neural network forecasting of the British Pound/US Dollar exchange rate. *Omega, The International Journal of Management Science*, 26(4), pp.495–506.
- Zhang, G.P., 2003a. Business forecasting with artificial neural networks: An overview. In G. P. Zhang, ed. *Neural Networks in Business Forecasting*. Idea Group Publishing, p. 296.
- Zhang, G.P., 2003b. Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, pp.159–175.
- Zhang, G.P., Patuwo, B.E. & Hu, M.Y., 1998. Forecasting with artificial neural networks : The state of the art. *International Journal of Forecasting*, 14, pp.35–62.

8 Apêndices

Apêndice A – Medidas de Performance estatística

Erro quadrático médio	$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (\hat{y}_i - y_i)^2}$
Erro médio percentual absoluto simétrico	$SMAPE = \frac{\sum_{i=1}^T \hat{y}_i - y_i }{\sum_{i=1}^T (y_i + \hat{y}_i)}$
Porcentagem Correta de Mudança de Direção	$CDC = \frac{100}{T} \sum_{i=1}^T D_i \text{ onde, } \begin{cases} D_i = 1 \text{ se } y_i \times \hat{y}_i > 0 \\ D_i = 0 \text{ se } y_i \times \hat{y}_i \leq 0 \end{cases}$

Nota: y_i corresponde ao valor no momento i , \hat{y}_i ao valor previsto, e T corresponde ao período considerado

Apêndice B -- Valores p dentro da amostra para teste BDS sobre resíduos semanais e mensais dos modelos individuais, antes de hibridação

m	EUR/USD Semanal		JPY/USD Semanal		EUR/USD Mensal		JPY/USD Mensal	
	RNA	ARIMA	RNA	ARIMA	RNA	ARIMA	RNA	ARIMA
	[8,14,7,1]	(2,1,5)	[7,16,5,1]	(0,1,0)	[5,11,6,1]	(1,1,7)	[7,16,5,1]	(2,1,3)
2	0.124	0.0594	~0	~0	0.192	0.875	0.099	0.739
3	0.036	0.006	~0	~0	0.126	0.408	0.031	0.511
4	0.016	0.008	~0	~0	0.094	0.312	0.020	0.411
5	0.009	0.001	~0	~0	0.087	0.133	0.011	0.310
6	0.002	0	~0	~0	0.079	0.073	0.003	0.137
7	0	0	~0	~0	0.052	0.088	~0	0.098
8	0	0	~0	~0	0.030	0.079	~0	0.089

Apêndice C -- Valores dentro da amostra do teste LWG para os erros dos modelos individuais

	EUR/USD Semanal		JPY/USD Semanal		EUR/USD Mensal		JPY/USD Mensal	
	RNA	ARIMA	RNA	ARIMA	RNA	ARIMA	RNA	ARIMA
	[8,14,7,1]	(2,1,5)	[7,16,5,1]	(0,1,0)	[5,11,6,1]	(1,1,7)	[7,16,5,1]	(0,1,0)
LWG	0	3	2	788	0	1	0	0

Nota: O teste LWG corresponde ao número de RNA's aleatórias no teste que apresentaram capacidade de modelar não linearidades entre os resíduos e a variável dependente, significativas a 5%. O teste LWG inicializa 1000 RNA's com pesos aleatórios. Quando existem mais de 50 significativas, rejeita-se H0, de que não existem não linearidades por modelar, a um nível de significância de 5%.

Apêndice D – Variáveis de entrada selecionadas pelo algoritmo genético e valor de α para cada RNA multivariada no estudo

EUR/USD Semanal	JPY/USD Semanal	EUR/USD Mensal	JPY/USD Mensal
RNA [8,14,7,1]	RNA [7,16,5,1]	RNA [5-11-6-1]	RNA [7,16,5,1]
Eur/Usd (1,2)	Jpy/Usd (1)	Eur/Usd (1,2)	Jpy/Usd (1,2)
Libor US 1M (1,2)	Libor US 1M (1,2)	Libor US 1M (1)	Libor US 1M (1,2)
Preço Petróleo (1)	Preço Petróleo (1)	Preço Petróleo (1)	Preço Petróleo (1)
Preço Ouro (2)	Libor JP 1M (1,2)	IPC US (1)	IPC US (1)
SP500 (1)	Libor JP 3M (1)		Desemprego US (1)
Euribor 1M (2)			
$\alpha = 0.84$	$\alpha = 0.80$	$\alpha = 0.05$	$\alpha = 0.025$

Nota: Devido a α , não existe garantia de que todas as variáveis selecionadas pelo AG e transcritas na tabela são utilizadas para realizar previsão. O valor de α pode, durante a estimação, forçar a zero os pesos sinápticos associados a algumas variáveis de entrada, e a alguns nós escondidos, fazendo com que para certos períodos de previsão estes não sejam utilizados na estimação.

Entre parêntesis curvos encontram-se os desfasamentos selecionados das variáveis.

Apêndice E - Sumarização de alguns estudos efetuados para a previsão de taxas de câmbio utilizando Redes Neurais Artificiais

Ano	Autores	Modelos		Base Dados	Taxas de câmbio	Período temporal	Resultados
		Tipologia	Híbrido Multivariado				
2011	Khashei, Bijari	Arima – RNA de retropropagação Algoritmos de Corte	Sim	Não	1994-2003Inr/Usd	Semanal (h=1)	A abordagem híbrida proposta produz melhores resultados que uma RNA univariada, um modelo linear autorregressivo e que o modelo passeio aleatório fora da amostra
2011	Koloch, Rubaszek, Skrzypczyńsk	RNA de retropropagação	Não	Não	1999-2009Pln/Eur Pln/Usd Pln/Gbp Pln/Chf Pln/Czk	Semanal h=1,4,8,12,26,52	Embora para alguns horizontes e alguns pares os moedas a RNA tenham apresentado resultados melhores que um passeio aleatório, no global conclui-se que o modelo passeio aleatório apresenta melhores resultados
2010	Pacelli, Bevilacqua, Azzollini	Algoritmo Genético com RNA de retropropagação	Não	Sim	1999-2009Eur/Usd	Diário (h=1)	Encontra evidência de que a taxa de câmbio não evolui de forma totalmente aleatória, e assim, em parte previsível
2007	Panda, Narasimhan	RNA retropropagação	Não	Não	1994-2003Inr/Usd	Semanal (h=1)	A RNA apresenta melhores resultados que um modelo de passeio aleatório fora da amostra.
2006	Seliem	RNA retropropagação	Não	Sim*	----- Eur/Usd Gbp/Usd Jpy/Usd Usd/Chf	Diário, Semanal Mensal (h=1)	Com dados mensais a RNA revela baixa capacidade de de previsão, contrariamente aos restantes.

Apêndice E - Continuação

Ano	Autores	Modelos	Base		Taxas de câmbio	Período temporal	Resultados
		Tipologia	Híbrido	Multivariado			
2003	Zhang	Arima-RNA	Sim	Não	1980-1992Gbp/Usd	Semanal (h=1)	Conclui-o que para a taxa de câmbio considerada o modelo de passeio aleatório apresentava resultados inferiores ao modelo híbrido univariado.
2003	Kamruzzaman, Sarker	RNA (testa 3 algoritmos diferentes)	Não	Não	1991-2002Aud/Gbp Aud/Jpy Aud/Usd Aud/Chf Aud/Sgd Aud/Nzd	Semanal (h=1)	Os 3 algoritmos de treino diferentes produzem boas previsões, mas o algoritmo tradicional de retropropagação é o pior.
2002	Dunnis, Williams	Forward Stepwise com RNA de retropropagação	Não	Sim	1994-2001Eur/Usd	Diário (h=1)	Em termos estatísticos o modelo não apresenta resultados claramente melhores aos modelos lineares. Em termos financeiros este é melhor.
2000	Gradojevic, Yang	RNA de retropropagação	Não	Sim	1990-2000Cad/Usd	Diário (h=1, 7)	As duas RNA selecionadas pelos autores comportam-se melhor fora da amostra que modelos lineares e de passeio aleatório.
1998	Jamal, Sundar	RNA de retropropagação	Não	Não	1984-1993Usd/Dem Usd/Frf	Mensal (h=1, 2, 3, 4)	A RNA é superior na previsão fora da amostra face a um modelo multivariado de regressão linear

* Utiliza unicamente médias móveis como dados de entrada, assim como o valor mais alto e mais baixo negociado em períodos anteriores. Utiliza ainda os valores a que se negociava a moeda ao encerramento de cada período.

Nota: “h” representa o numero de períodos à frente para os quais os autores realizam previsões.

